فصلنامه صنعت بیمه / سال پانزدهم / شماره ۶۰ / زمستان ۱۳۷۹ / صفحات ۴۷تا۶۲

مدل سازی ادعاهای خسارت‌های بزرگ و کاربرد آن در بیمه آتش‌سوزی شرکت‌های بیمه

حبیب اسمعیلی جندابه[۱]

چکیده

در این مقاله ضمن بررسی اجمالی نظریه مخاطره و معادله اساسی آن، به بیان رفتار حدی مجموع متغیرهای تصادفی و ماکسیمم متغیرها پرداخته و سپس روش‌های آماری برای بررسی رفتار تصادفی ماکسیمم‌ها را بیان می‌کنیم. در پایان این روش‌ها را روی داده‌های واقعی شرکت بیمه در بخش آتش سوزی اجرا و سپس به برخی پرسشها در این زمینه پاسخ می‌دهیم.

واژگان کلیدی

ادعاهای خسارت‌های بزرگ، نظریه مخاطره، کرانگین، دامنه جذب، تابع میانگین مازاد.

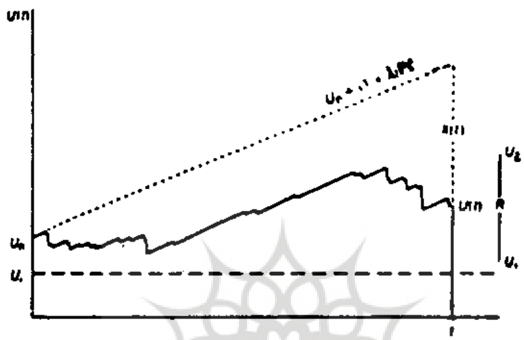
مقدمه

شرکت بیمه‌ای را در نظر می‌گیریم که با سرمایه اولیه u_0 کار خود را آغاز و با گذشت زمان مطابق تابع $p(t)$ از مشتریان خود حق بیمه دریافت می‌کند. در زمان‌های تصادفی T_1،T_2،... مدعیان خسارت به شرکت بیمه مراجعه و خسارت‌های خود را با اندازه‌های تصادفی X_1،X_2،... از شرکت مطالبه می‌کنند. ذخیره مخاطراتی شرکت بیمه یا سرمایه در دسترس شرکت در زمان t از معادله زیر محاسبه می‌شود. $U(t) = u_0 + p(t) - S(t)$

---

۱. نویسنده از دانشکده علوم ریاضی دانشگاه شهید بهشتی این مقاله را بر اساس پایان نامهٔ کارشناسی ارشد خود نوشته است (استاد راهنما: دکتر محمدرضا مشکانی، استاد مشاور: دکتر ابوالقاسم وحیدی اصل)

که در آن $S(t)$ مجموع خسارت‌های ادعا شده تا زمان t است $\left[S(t)=\sum_{i=1}^{N(t)} X_i\right]$ تابع S(t) به صورت مجموع تعدادی تصادفی از متغیرهای تصادفی تعریف می‌شود که جزء اصلی و تصادفی معادله مخاطره را تشکیل می‌دهد. تابع P(t) یک تابع قطعی است که معمولاً به صورت خطی از t (ct) تعریف می‌شود اغلب برای اعمال هزینه‌های جاری شرکت این تابع را در ضریب $(1+\lambda)$ نیز ضرب می‌کنند که λ را سربار ایمنی [1] گویند. شکل زیر نمایش فرایند ذخیره مخاطراتی شرکت بیمه است.



u_2 مقداری است که اگر ذخیره مخاطراتی شرکت بیمه از آن حد کمتر شود، شرکت قادر به برآورد تعهداتش نیست و در این حالت آن را ورشکسته خوانند. چگونگی مدل سازی فرایند ذخیره مخاطراتی از مباحث مربوط به نظریه مخاطره است که از بحث آن خودداری می‌کنیم. فقط ذکر این نکته الزامی است که مدل مورد بحث در نظریه مخاطره در حالت کلی مدل تجدید و در حالت خاص‌تر مدل کرامر ـ لاندبرگ است. مدل کرامر ـ لاندبرگ مدلی است که در آن فرض می‌شود، ادعاهای مراجعه کننده به شرکت دارای مقادیر تصادفی، مثبت و مستقل با تابع توزیع مشترک F، میانگین متناهی μ و واریانس متناهی σ^2 است. همچنین ادعاها در زمان‌های تصادفی رخ داده و زمان‌های بین ادعاها متغیرهای تصادفی iid با تابع توزیع نمایی باشند. با بیان این مطلب به طور اجمالی در مورد رفتار S(t) بحث می‌کنیم.

۱. نوسانهای مجموع متغیرهای تصادفی

فرض کنیم S_n مجموع n متغیر تصادفی iid باشد. می‌خواهیم ببینیم که اگر S_n به طور

1.safty loading

مناسب مرکزی و استاندارد شود به چه قانون حدی[1] همگراست.

تعریف ۱ـ۱ فرض کنید X, X_1, X_2 متغیرهای تصادفی iid باشند. متغیر تصادفی X یا توزیع آن را پایدار نامیم، اگر در رابطه زیر به ازای اعداد نامنفی c_1, c_2 و اعداد حقیقی مناسب $a(c_1, c_2), b(c_1, c_2)$ صدق کند:

$$c_1 x_1 + c_2 x_2 \underline{\underline{d}}\ b(c_1, c_2)\, x + a(c_1, c_2)$$

اگر S_n مجموع متغیرهای تصادفی پایدار باشد، با استفاده از رابطه فوق به ازای بعضی ثابت‌های a_n, b_n داریم:

$$\frac{S_n - a_n}{b_n} \underline{\underline{d}}\, X$$

در واقع حالت فوق فقط برای توزیع‌های پایدار برقرار است. مثلاً اگر توزیع خسارت‌ها نرمال باشد، توزیع مجموع خسارت‌ها نیز نرمال خواهد بود. زیرا توزیع نرمال یک توزیع پایدار است. توزیع‌های ناتباهیده پایدار، رده‌ای از توزیع‌ها را تشکیل می‌دهند که این رده با رده قوانین حدی ناتباهیده ممکن برای مجموع‌های متغیرهای تصادفی iid بر هم منطبق است. در حالت کلی مجموع مرکزی شده و استاندارد شده S_n به طور ضعیف به توزیع پایدار G_α می‌گراید:

$$\frac{S_n - a_n}{b_n} \xrightarrow{d} G_\alpha$$

در این صورت گوییم متغیر تصادفی X یا توزیع آن به دامنه جذب توزیع پایدار G_α تعلق دارد. مهم‌ترین قضیه در این مورد، قضیه حدی مرکزی است که چنین بیان می‌کند: تمامی توزیع هایی که حداقل دارای گشتاور مرتبه دوم متناهی هستند در دامنه جذب توزیع نرمال قرار دارند. بنابراین توزیع حدی S_n که به طور مناسب استاندارد شده است، در اغلب موارد نرمال است. مطالب زیادی در این مورد وجود دارد که بیان می‌کند حتی زمانی که n نیز تصادفی است باز نیز در اغلب موارد توزیع حدی متغیر استاندارد شده S_n، نرمال استاندارد است (در مأخذ ۳ مطالب بیشتری آمده است).

---

1. limit law

۲. نوسانهای ماکسیمم‌ها

ابتدا واژه «کرانگین»(۱) را بررسی می‌کنیم. اگر یک رشته از اعداد را به ترتیب افزایشی یا کاهشی مرتب کنیم داده‌هایی که در دو سر این دنباله قرار می‌گیرند، یعنی کوچک‌ترین و بزرگ‌ترین داده‌ها را کرانگین‌های این مجموعه می‌نامیم. در بعضی از شاخه‌های علوم نظیر هیدرولوژی هر دو مقادیر برای ما مهم‌اند، زیرا میزان بارش اندک به خشک‌سالی و میزان بارش زیاد به پیدایش سیل منجر می‌شود. اما در اغلب موارد از جمله در بیمه کرانگین‌های بالایی یا ماکسیمم‌ها برای ما اهمیت دارند. در ادامه منظور از کرانگین‌ها، ماکسیمم‌های نمونه‌های تصادفی است. اگر $X_1, X_2, \ldots, X_n$ دنباله‌ای از متغیرهای تصادفی iid با تابع مشترک F باشند، ماکسیمم نمونه را با M_n نشان می‌دهیم:

$$M_n = max\,(X_1, \ldots, X_n)$$

توزیع دقیق M_n عبارت است از:

$$P(M_n \le x) = P(X_1 \le x, X_2 \le x, \ldots, X_n \le x) = F^n(x)$$

کرانگین‌ها نزدیک به انتهای تکیه‌گاه توزیع رخ می‌دهند، بنابراین رفتار مجانبی M_n با دم توزیع F نزدیک به نقطه انتهایی مرتبط است. اگر نقطه انتهایی سمت راست را با x_F نشان دهیم، در این صورت وقتی $n \to \infty$ آن‌گاه $M_n \xrightarrow{p} x_F$. مهم‌ترین قضیه‌ای که برای ماکسیمم‌ها مطرح می‌شود، قضیه فیشر-تیپو است. اهمیتی که این قضیه برای کرانگین‌ها دارد، به اندازه اهمیت قضیه حدی مرکزی برای مجموع است. در واقع این قضیه توزیع مجانبی ماکسیمم‌های استاندارد شده را ارائه می‌دهد.

قضیه فیشر - تیپو: فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی iid باشند. اگر ثابت‌های نرم‌گر $c_n > 0$، $d_n \in R$ و تابع توزیع ناتباهیده H وجود داشته باشند به طوری که:

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H$$

---

1. extreme

آن گاه H متعلق به یکی از توابع زیر است:

الف) فره شه
$$\Phi_\alpha = \begin{cases} \cdot & ,X \leq \cdot \\ e^{-x^{-\alpha}} & ,x > \cdot \end{cases}$$

ب) وایبول
$$\Psi_a(x) = \begin{cases} e^{-(-x)^a} & x \leq \cdot \\ 1 & x > \cdot \end{cases} \quad , a > \cdot$$

پ) گامبل
$$\Lambda(x) = e^{-x} \qquad x \in R$$

متناظر با توزیع‌های پایدار برای مجموع، توزیع‌های ماکس-پایدار[1] را برای ماکسیمم‌ها خواهیم داشت. همچنین متناظر با دامنهٔ جذب برای مجموع، دامنه جذب ماکسیمم [2] رابرای ماکسیمم‌های نمونه‌های تصادفی داریم، به طور کلی متغیر تصادفی X یا توزیع آن بـه دامـنه جـذب مـاکسیمم تـوزیع H مـتعلق است اگـر و فـقط اگـر ثـابت‌های $d_n \in R, c_n > \cdot$ وجود داشته باشند، به طوری که:

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H$$

توزیعهای مختلف با خصوصیات و رفتارهای دمی متفاوت به دامنه‌های جذب ماکسیمم هر یک از توزیعهای سه گانه مقدار کرانگین متعلق هستند. در زیر به طور خلاصه به این دامنه جذب ماکسیمم اشاره می‌کنیم:

الف) دامنهٔ جذب ماکسیمم توزیع فره شه: توزیع F به دامنه جذب ماکسیمم Ψ_a متعلق است اگر و فقط اگر تابع دمی آن را بتوان به صورت زیر نوشت:

$$\overline{F}(x) = x_{\cdot}^{-a} L(x)$$

که در آن L یک تابع کند تغییر[3] است یعنی به ازای هر t>0

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1$$

1. max stable 2. maximum domain of attraction

3. Slowly Varying

در این صورت می‌نویسیم: $F \in MDA(\Psi_a)$.

ب) دامنهٔ جذب ماکسیمم توزیع وایبول: توزیع F به دامنهٔ جذب ماکسیمم توزیع Ψ_a متعلق است اگر و فقط اگر $x_F < \infty$ و تابعی کند تغییر نظیر L موجود باشد به طوری که

$$\overline{F}\left(x_F - \frac{1}{x}\right) = x^{-a} L(x)$$

پ) دامنهٔ جذب ماکسیمم توزیع گامبل: اگر توزیع F یک تابع فون میزس [1] باشد آن گاه در دامنه جذب ماکسیمم توزیع گامبل قرار دارد ($F \in MDA(\Lambda)$).

یادآوری می‌کنیم که تابع F را یک تابع فون میزس با تابع کمکی a(.) نامیم اگر F را بتوان به شکل زیر نمایش داد:

$$\overline{F}(x) = c.exp\left\{-\int_z^x \frac{1}{a(t)}\,dt\right\} \quad ,z < x < x_F$$

معمولاً برای نمایش توزیعهای مقدار کرانگین، از یک شکل کلی به نام توزیع مقدار کرانگین تعمیم یافته که توسط جنکینسن [2] و فون میزس ارائه شده است استفاده می‌شود. در این حالت یک پارامتر اضافی که به پارامتر شکل [3] معروف است، وارد توزیع می‌شود و بر حسب این که این پارامتر مثبت، منفی یا صفر باشد هر سه توزیع مقدار کرانگین فره شه، وایبول و گامبل نتیجه می‌شود. نمایش جنکینسن ـ فون میزس توزیعهای مقدار کرانگین به صورت زیر است:

$$H_\xi(x) = \begin{cases} e^{-(1+\alpha)^{\frac{-1}{\xi}}} & , \ \xi \neq 0 \\ e^{-e^{-x}} & , \ \xi = 0 \end{cases}$$

برای $1 + \xi x > 0$. بنابراین اگر $\xi > 0$ آن گاه $x > -\frac{1}{\xi}$ اگر $\xi < 0$ آنگاه $x < \frac{-1}{\xi}$ اگر $\xi = 0$

1. Von Misses 2. Jenkinson 3. shape parameter

آنگاه $x \in R$ (برای جزئیات بیشتر در این مورد و تعیین ثابتهای نرم‌گر می‌توان به مأخذ ۳ مراجعه کرد).

گاهی در مدل سازی پیشامدها، ماکسیمم‌های دوره‌ای مدنظر نیستند، بلکه ماکسیمم‌های فراتر از یک مقدار آستانه‌ای مانند u مد نظرند. مثلاً تمامی خسارتهایی که مبلغ خسارت از مقدار u بیشتر شود یا تمامی بارشهایی که میزان بارش از مقدار آستانه‌ای u فراتر رود، ممکن است حادثه ساز باشند. برای مدل سازی این گونه پیشامدها معمولاً از مدل توزیع پارتوی تعمیم یافته (GPD) که به صورت زیر تعریف می‌شود استفاده می‌کنند:

$$G_\xi(x) = \begin{cases} 1-(1+\xi x)^{\frac{-1}{\xi}} & ,\xi \neq 0 \\ 1-e^{-x} & ,\xi = 0 \end{cases}$$

با جانشین کردن $\frac{x-\mu}{\sigma}$ به جای x خانواده مکان ـ مقیاس $G_{,\mu,\sigma}$ حاصل می‌شود، لکن در عمل معمولا از $G_{\xi,\beta}$ به شکل زیر استفاده می‌شود:

$$G_{\xi,\beta} = 1-(1+\xi\frac{x}{\beta})^{\frac{-1}{\xi}} \quad , x \in D(\xi,\beta)$$

۳. روش‌های آماری برای پیشامدهای کرانگین

در این بخش هدف، ارائه بعضی از راهکارهای گرافیکی برای تشخیص و مدل سازی کرانگین‌ها در مجموعه داده‌های واقعی است. همچنین در مورد روش‌های برآورد پارامترهای توزیع مقدار کرانگین بحث می‌شود.

۱.۳ تحلیل اکتشافی داده‌ها برای کرانگین‌ها

نمودارهایی وجود دارد که از طریق آن‌ها می‌توان به خصوصیات توزیع از جمله رفتار

دمی آن پی برد که این موضوع می‌تواند به مدل سازی صحیح داده‌ها کمک کند. برخی از این روش‌های گرافیکی را در زیر ارائه خواهیم داد.

۱.۱.۳ نمودارهای احتمالی و چندکی

فرض می‌کنیم $X_{1,n} \le \dots \le X_{n-1,n} \le X_{n,n}$ آماره‌های مرتب نمونه تصادفی iid از توزیع F باشد. اگر F پیوسته باشد در این صورت $U_i = F(X_i) \quad , \quad (i = 1, \dots, n)$ نمونه تصادفی iid از توزیع یکنواخت (۱و۰) است به علاوه

$$E\left[F(X_{k,n})\right] = \frac{n-k+1}{n+1} \qquad , k = 1,2,\dots,n$$

,

بر این اساس، نمودار احتمالی به صورت

$$\left\{ \left(F(x_{k,n}), \frac{n-k+1}{n+1}\right) \qquad , k = 1,2,\dots,n \right\}$$

تعریف می‌شود. یا می‌توان با تعریف وارون تعمیم یافته F به صورت $\overleftarrow{F}(t) = \inf\left\{x \in R : F(x) \ge t\right\}$ نمودار چندکی[1] را به صورت زیر تعریف کرد:

$$\left\{ \left(x_{k,n} . F^{\leftarrow}\left(\frac{n-k+1}{n+1}\right)\right) \;:\; K = 1,2,\dots,n \right\}$$

---

1. quantile function

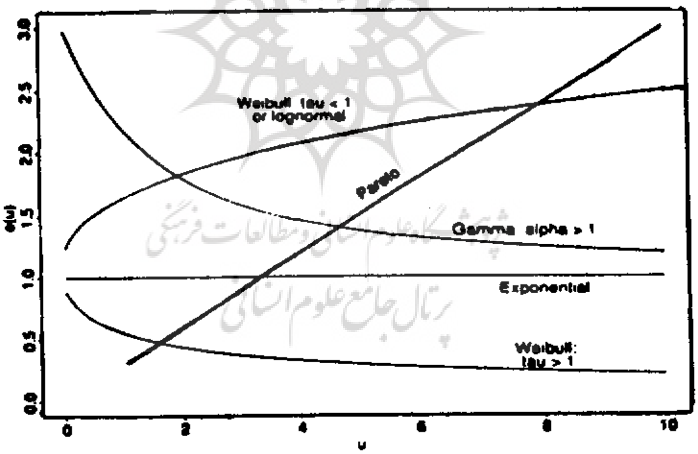
اگر داده‌ها دقیقا از توزیع F استخراج شده باشند، این نمودارها به صورت خطی دیده می‌شوند. بر حسب این که توزیع داده‌ها نسبت به توزیع مفروض F دم کلفت‌تر و یا دم نازک‌تر باشد، تقعر منحنی به سمت بالا یا پایین قرار می‌گیرد و هر چه ضخامت دم توزیع بیشتر یا کمتر شود تقعر منحنی نیز کمتر یا بیشتر می‌شود.

۲.۱.۳ تابع میانگین مازاد[1]

این تابع به شکل زیر تعریف می‌شود:

$$e(u) = E(X - u \mid X > u) \quad , \quad 0 \le u \le x_F$$

توابع توزیع با رفتار دمی متفاوت دارای توابع میانگین مازاد با شکل‌های گوناگون هستند. بنابراین از شکل نمودار میانگین مازاد تجربی می‌توان به رفتار دمی توزیع پی برد. در زیر نمودار توابع میانگین مازاد برای برخی از توزیعها رسم شده است.



۳.۱.۳ دوره بازگشت

دوره بازگشت در واقع میانگین زمان مورد انتظار برای رخداد یک پیشامد کرانگین است. اگر یک مقدار آستانه‌ای u را در نظر بگیریم، به سادگی می‌توان نشان داد که متغیر

---

1. the mean excess function

تصادفی زمان اولین مازاد از مقدار آستانه‌ای u، دارای توزیع هندسی با احتمال موفقیت $P = \overline{F}(u)$ است بنابراین دوره بازگشت برابر $\frac{1}{\overline{F}(u)} = \frac{1}{P}$ است. پس با مشخص شدن توزیع F می‌توان دوره بازگشت پیشامدهای فراتر از مقدار آستانه‌ای u را محاسبه کرد.

۲.۳ برآورد پارامترهای توزیع مقدار کرانگین

فرض کنیم $X_1, \dots, X_{n-1}, X_n$ نمونه تصادفی iid از توزیع مقدار کرانگین H_Θ باشد، در این صورت برای برآورد ماکسیمم درستنمایی، معادلات درستنمایی، فرم بسته‌ای ندارند. لذا حل این معادلات فقط به روشهای عددی امکان‌پذیر است. برآوردهای درستنمایی ماکسیمم برای حالت $\xi > -\frac{1}{2}$ خواص کلاسیک و خوب برآوردهای ML را دارا هستند ولی برای حالت $\xi \le -\frac{1}{2}$ این خواص را ندارند و پیشنهاد می‌شود از روشهای دیگری استفاده شود، از جمله برآوردهای خوبی که در این حالت‌ها پیشنهاد می‌شود برآوردهایی است که مبتنی بر روش گشتاورهای احتمالی ـ موزون است. در این روش گشتاورهای احتمالی ـ موزون نظری $w_r(\theta) = E(XH_\theta^r(X))$ را باگشتاورهای نمونه‌ای که به صورت $W_r(\Theta) = \frac{1}{n}\sum_{j=1}^{n} X_{j,n} H_\Theta^r(X_{j,n})$ محاسبه می‌شوند، برابر قرار می‌دهند و برآوردهای مورد نظر را محاسبه می‌کنند.

اگر بخواهیم برآورد پارامترهای مقدار کرانگین را به دست آوریم در حالی که نمونه مورد نظر از توزیع مقدار کرانگین H_θ استخراج نشده باشد، بلکه نمونه‌ای از توزیع F باشد که این توزیع در دامنه جذب ماکسیمم توزیع مقدار کرانگین H_θ قرار دارد. به عبارت دیگر

$$X_1, X_2, \dots, X_n \overset{iid}{\sim} F \quad , \quad F \in MDA(H_\theta)$$

در این حالت نیز برآوردگرهایی برای مهم‌ترین پارامتر توزیع یعنی پارامتر شکل ξ ارائه شده است که از جمله می‌توان به برآوردگر پیکاندار[1]، برآوردگر هیل[2] و برآوردگر

1. Pickand's estimator
2. Hill's estimator

DEdH[۱] اشاره کرد.

(برای توضیحات بیشتر در مورد برآوردگرها می‌توانید مأخذ ۳ را بخوانید).

۳.۳ آزمون نیکویی برازش

معمولاً بعد از تشخیص مدل و برآورد پارامترها، بررسی خوبی مدل مدنظر قرار می‌گیرد و این پرسش مطرح می‌شود که این مدل تا چه حدی می‌تواند برازنده داده‌های ما باشد. روشهای متفاوتی برای بررسی خوبی مدل وجود دارد که از آن جمله می‌توان به آزمونهای مجذور خی و کلموگروف ـ اسمیرنوف اشاره کرد. در مواقعی که دم توزیع برای ما اهمیت دارد، پیشنهاد می‌شود که از آزمون آندرسون ـ دارلینگ که آماره این آزمون عضوی از خانواده کرامر ـ فون میزس است، استفاده شود. خانواده کرامر فون ـ میزس یک خانواده از آماره هایی است که تفاوت بین $F_n(x)$, $F(x)$ را به صورت زیر اندازه‌گیری می‌کند:

$$W=\int_{-\infty}^{\infty}\left\{F_n(x)-F(x)\right\}^{2}\Psi(x)\,dx$$

که در آن تابع وزنی $\Psi(x)$ برای آزمون آندرسون ـ دارلینگ به صورت زیر تعریف می‌شود:

$$\Psi(x)=\frac{1}{F_\theta(x)\left(1-F_\theta(x)\right)}$$

علت انتخاب این آزمون برای توزیعهای مقدار کرانگین این است که این آماره بیشتر از سایر آماره‌ها به مشاهدات واقع در دم توزیع اهمیت می‌دهد.

۴. تحلیل ادعاهای خسارتهای بزرگ در بیمه آتش سوزی

در این بخش چگونگی جمع آوری و تحلیل داده‌های بیمه آتش سوزی یک شرکت بیمه را با استفاده از مطالب نظری فصلهای گذشته، بیان خواهیم کرد.

1. Deckers - Einmaho - deHaan

۱.۴ توصیف نوع اطلاعات و نحوه جمع آوری آنها

داده‌های مورد استفاده، داده‌های مربوط به بیمه آتش سوزی شرکت بیمه است که از بردروهای بخش آتش سوزی آن شرکت با درج تاریخ پرداخت خسارت طی دورهٔ زمانی مورد نظر، استخراج دور یک فایل ثبت شده است. دورهٔ زمانی مورد نظر به مدت ۵۲ ماه از ابتدای فروردین ماه ۷۴ تا پایان تیرماه ۷۸ بوده است. چون ماکسیمم‌های سالانهٔ فقط ۴ مشاهده بودند، از ماکسیمم‌های ماهانه به تعداد ۵۲ مشاهده برای تجزیه و تحلیل استفاده شده است. این کار همواره امکان‌پذیر نیست. مثلاً در هیدرولوژی، داده‌های ماهانه به سبب تغییر میزان بارش در فصلهای مختلف سال همبسته می‌شوند و نمی‌توان از روشهای بیان شده که مبتنی بر استقلال مشاهدات ماکسیمم از یکدیگر است در دوره‌های ماهانه یا فصلی، استفاده کرد. لکن آتش سوزی‌ها و به خصوص آتش سوزی‌های صنعتی (که ادعای خسارتهای بزرگ اغلب به آن مربوط می‌شود). در فصول مختلف سال رخ می‌دهند و مستقل از ماههای سال هستند.

۲.۴ تجزیه و تحلیل اطلاعات

در مرحله تجزیه و تحلیل اطلاعات متوجه اشکالی در داده‌ها شدیم که از تورم ناشی می‌شد. بدین ترتیب که خسارتهای رخ داده در سال ۱۳۷۴ با خسارتهای رخ داده در سال ۱۳۷۷ یا ۱۳۷۸ هم ارزش نبودند و به این دلیل مبالغ ادعا شده تنها ناشی از عمق خسارت نبوده بلکه تورم نیز در میزان آن دخالت داشته است. برای رفع این مشکل باید اثر تورم را از داده‌ها خارج و آنها را در یک زمان خاص تجزیه و تحلیل کرد. برای این منظور با استفاده از شاخص کل بهای عمده فروشی کالاها، نرخ تورم از داده‌ها خارج و تمامی داده‌ها به یک زمان آورده شدند. علل استفاده از این شاخص عبارت است از:

الف) چون ادعاهای خسارتهای بزرگ بیشتر در بخش مشاغل رخ داده بودند، بنابراین استفاده از شاخص مصرفی کالاها و خدمات خانوار که بر اساس میزان کالاها و خدمات مصرفی خانوار محاسبه می‌شود صحیح نیست. به علاوه در محاسبه شاخص اخیر خدمات نیز وزنی را شامل شده‌اند که خدمات نظیر خدمات پزشکی و ... بر اثر

آتش سوزی از بین نمی‌روند.

ب) شاخص کل بهای عمده فروشی کالاها علاوه بر این که شامل خدمات نیست براساس میزان کالاهایی است که در بازار و مشاغل در جریان است.

پ) شاخصهای ضمنی در بخشهای مختلف اقتصادی فقط تا سال ۱۳۷۶ موجود بودند و برای سال ۱۳۷۷ نیز به صورت مقدماتی محاسبه شده بود. به علاوه برای استفاده از شاخصهای ضمنی باید تک تک خسارتها به تفکیک بررسی شوند. آن بخش از خسارت که به ساختمان مربوط می‌شود با استفاده از شاخص مربوط به خود و ... باید تورم زدایی شوند. لکن در اغلب موارد که خسارتها در بردروها نوشته می‌شوند این تفکیکات صورت نمی‌گیرد.

با این اوصاف، بهترین شاخص برای تورم زدایی شاخص کل بهای عمده فروشی کالاها تشخیص داده شد. چون در بخش سوم ثابت کردیم که توزیع مشاهدات ماکسیمم، توزیع مقدار کرنگین تعمیم یافته است، لذا پارامترهای توزیع مقدار کرانگین تعمیم یافته را به روش درستنمایی ماکسیمم و براساس مشاهدات برآورد کردیم که نتایج زیر حاصل شد:

$$\hat{\xi} = -۰/۱۴۱۷۶۷۲$$
$$\hat{\sigma} = ۵۰۳۲۳۰/۳$$
$$\hat{\mu} = ۳۲۸۳۲/۴۹$$

در این حالت معمولاً اولین فرضی که مورد آزمون قرار می‌گیرد، در رابطهٔ با صفر بودن پارامتر شکل است یعنی: $H_۰ : \xi = ۰$

خطای استاندارد پارامتر شکل را نیز برآورد کردیم، داریم: $se(\hat{\xi}) = ۰/۰۲۸۶۲۵۲۶$

با آزمون فرض فوق $H_۰$ رابه طور قوی رد می‌کنیم، با توجه به منفی بودن پارامتر شکل، توزیع وایبول به عنوان توزیع مقدار کرانگین مورد نظر، حاصل می‌شود. بنابراین مدل وایبول با شرایط ارائه شده به داده‌های ماکسیمم برازنده می‌شود. سپس با استفاده از

آزمون مجذور خی و کلموگروف ـ اسمیرنوف آزمون نیکویی برازش را اجرا کردیم که نتیجه، در هر دو حالت برازندگی مدل وایبول را تأیید کرد. توزیع وایبول حاصل شده عبارت است از:

$$H(x)=exp\left\{-(۱-۰/۱۴۱۷۶۳۲\frac{x-۳۲۸۳۲}{۵۰۳۲۳۰/۳})^{۷/۰۵۴}\right\}=\Psi_{\alpha}(-(۱-\frac{x-۳۲۸۳۲/۴۹}{۳۵۴۹۷۸۶/۵۴}))$$

که در آن Ψ نمایانگر توزیع وایبول با شاخص $\alpha=\frac{-۱}{\xi}=۷/۰۵۴$ است. در این جا متذکر می‌شویم که مدل فوق بر پایهٔ قیمتهای پایهٔ سال ۱۳۷۴ و به هزار ریال است. اکنون با مشخص شدن توزیع، پرسشهایی مطرح و به آنها پاسخ می‌دهیم.

فرض کنیم مدیریت بخش آتش سوزی شرکت بیمه در نظر دارد دورهٔ بازگشت پیشامدهای فراتر از یک میلیارد ریال را به ارزش پول سال ۷۴ محاسبه کند. این مبلغ به ارزش پول رایج در مرداد ۱۳۷۸ معادل مبلغی بیش از دو میلیارد و دویست و پنجاه میلیون ریال است داریم:

$$p=\overline{H}(۱\ ۰۰۰۰۰)=۰/۱$$

بنابراین به طور متوسط هر دو ماه یک بار باید انتظار یک پیشامد فراتر دومیلیارد و دویست و پنجاه میلیون ریالی به ارزش پول مرداد ۱۳۷۸ را داشته باشیم. همین طور برای سایر پیشامدها نیز می‌توان دوره بازگشت را محاسبه کرد.

فرض کنیم مدیریت بخش آتش سوزی شرکت بیمه فوق مایل است دورهٔ بازگشت و مقدار آستانه‌ای u را طوری تعیین کند که حداقل طی ۲۰ ماه آینده بیشتر از ۵ درصد خطر شکست وجود نداشته باشد. اگر L(u) را زمان اولین مازاد از مقدار آستانه‌ای u تعریف کنیم، در این صورت L(u) دارای توزیع هندسی است و مسأله به این صورت مدل بندی می‌شود:

$$p(L(u) \leq ۲۰) \leq ۰/۰۵$$

شکست برای ماه iام را به صورت $\{X_i > u\}$ تعریف می‌کنیم بنابراین:

$$P(L(u) \leq ۲۰) = ۱ - (۱ - p)^{۲۰} = ۰/۰۵$$

با حل معادله فوق p=۰/۰۰۲۵۶۱۴ حاصل می‌شود بنابراین دورهٔ بازگشت E(L(u))= ۳۹۰ ماه است.

با محاسبهٔ حد آستانه‌ای مربوط داریم:

$$u_{۳۹۰} = F^{\leftarrow}(۱ - \frac{۱}{t}) = F^{\leftarrow}(۰/۹۹۷۴۳۶) = ۲۰۵۸۷۰۷$$

پس مدیریت بخش آتش سوزی برای چنین ریسکی لازم است در خصوص پیشامدهای ۳۲/۵ ساله با مقدار آستانه‌ای دو میلیارد و پنجاه و هشت میلیون ریال به پول رایج در فروردین ۱۳۷۴ برنامه‌ریزی کند.

همچنین اگر مدیریت مزبور بخواهد خسارتهای آتش سوزی فراتر از یک مقدار آستانه‌ای را مدل سازی کند، می‌تواند از توزیع پارتوی تعمیم یافته که در بخش سوم شرح داده شد، استفاده کند. برای مثال برای اتکایی کردن خسارتهای بیش از ده میلیون ریال (به پول رایج در فروردین ۱۳۷۴) مدل زیر حاصل می‌شود:

$$G_{\xi,\sigma,u}(x) = ۱ - (۱ + ۰/۹۶۷۱۹۸ \frac{x}{۱۵۸۲۳/۳۶})^{-۱/۰۳۳۹۱۴} \quad , \quad x \geq ۰$$

لازم به ذکر است که مدل فوق براساس مشاهدات آتش سوزی همان شرکت بیمه که طی سال ۱۳۷۵ رخ داده، تشکیل شده است. برای محاسبهٔ حق بیمه اتکایی و سایر

ملزومات برای اتکایی کردن، لازم است مشاهدات بیشتری طی دوره‌های طولانی‌تر جمع آوری شود و نباید فقط به مشاهدات یک سال اکتفا کرد. در پایان متذکر می‌شویم که برآوردهای فوق نیز به روش ماکسیمم درستنمایی محاسبه شده است. پیشنهاد می‌شود که اگر پارامتر شکل توزیع کوچک‌تر از $\frac{۱}{۲}$ - شد از برآوردهای گشتاوری احتمالی ـ موزون استفاده شود.

منابع

1. Beirlant. J. and Teugels, J.L. (1992) "Modelling large claim in non-life insurance" Insurance: Math. & Econcomi 11, 17-29.

2. Daykin, C. Pentikainen, T., Pesonen, M. "Practical Risk Theory for Actuaries" Chapman & Hall.

3. Embechts, P., Kluppelberg, C. Mikson, T. (1997) "Modelling Extremal Events for Insurance and Finance" Springer, New York.

4. Grandell, J. (1997) "Mixed Poisson Prrocese" Chapman & Hall, London.

5. Rootzen, H. and Tajvidi, N. (1997) "Extreme Value Statistics and wind storm losses: a case study. "Scan Actuar. J.PP 70-94.

۶. ضرغامی، سیما. بررسی آماره‌های فرین در رژیم بادهای خلیج فارس و دریای عمان، پایان نامهٔ کارشناسی ارشد، دانشگاه شهید بهشتی، ۱۳۷۱.

۷. یاری، غلامحسین. مطالعه آماری پیشامدهای طبیعی فرین براساس نظریهٔ مقادیر فرین مرتبط با آماره‌های ترتیبی، پایان نامهٔ کارشناسی ارشد، دانشگاه شهید بهشتی، ۱۳۶۵.